# Lamarck and Immunity : Somatic and Germline Evolution of Antibody Genes

E J Steele

C.Y.O'Connor ERADE Village Foundation
PO Box 5100, Canning Vale South,
Canning Vale, WA, 6155
✉ ejsteele@cyo.edu.au

*Manuscript received: November 2009 accepted: February 2010*

## Abstract

Current work on the mechanism of hypermutation of somatically rearranged antibody variable (V) genes shows that the most likely mechanism involves both direct DNA modification (deamination of cytosines to uracils by AID deaminase) and strand nicking plus mRNA editing (deamination of adenosine to inosine via the ADAR1 deaminase) coupled to a reverse transcription process to fix RNA sequence modifications in V gene DNA – most likely involving the repair enzyme DNA polymerase eta (η) known to be an efficient reverse transcriptase *in vitro*. The DNA sequence patterns of families of similar germline V genes reveals that many features of somatically mutated and antigen-selected variable genes appear written into the germline V gene arrays of the immune system. Lamarckian gene feedback and cellular reverse transcription, coupled to Darwinian antigen binding selection of somatically mutated V genes, are concepts which appear necessary for a more complete understanding how the V gene complex has evolved. Antibody variable (V) genes of the immune system have therefore been used to test ideas on reverse transcriptase-coupled soma-to-germline feedback in a complex multicellular system. Such feedback constitutes a violation of Weismann's Barrier and thus support for some type of Lamarckian gene feedback operative during the evolution of the vertebrate immune system.

**Keywords**: Lamarck, immunity, somatic and germline evolution, antibody genes; evolution

## Introduction

In this paper I will review the main findings of our studies on the mechanism of antigen-driven somatic hypermutation of rearranged antibody (immunoglobulin, Ig) variable region genes (so called VDJs) and the impact of this somatic genetic diversity on the germline V segment repertoire. I will then draw general conclusions on the origins of genome diversity. Many details of this work have already been covered in major reviews (Steele *et al.* 1993; Rothenfluh *et al.* 1995; Steele *et al.* 1997; Blanden *et al.* 1998) and in our 1998 book *Lamarck's Signature* (Steele *et al.* 1998). More recent work will be cited in the body of the text.

## General Concepts: Molecular-Cellular Immunology and Evolution

It is generally agreed that the primary evolutionary purpose of the immune system of vertebrates is the protection of the individual against disease. The proteins and carbohydrates which make up the cell walls, viral coats and secreted microbial toxins constitute the foreign antigens which individual immune systems need to react against to preserve the integrity of the body. The system consists of highly mobile blood white cells (lymphocytes) which come in two main categories, B cells and T cells which circulate from blood to lymph via a complex network of lymphatic vessels and capillaries. The complexity of the system almost rivals that of the brain and central nervous system (which in contrast consists of sessile or non-mobile nerve cells and fibres which generate their complexity in both their sheer cell numbers and cell-cell synaptic connections). The progenitors of the white cell lineages (B and T lymphocytes, monocytes, neutrophils, mast cells, polymorphonuclear leucocytes *etc.*) arise from stem cells in the bone marrow which produce many millions of hemopoietic cells on a daily basis (and, of course, the senescence of many other hemopoietic-derived cells as they exit the system). Hormonal and cytokine cell-to-cell communication no doubt allows the system to be co-ordinately controlled.

The primary evolutionary strategy of the immune system has been shaped by two selective forces, a) the requirement to respond to unexpected antigens thrown up by new infectious diseases, and b) the need to prevent autoimmune reactions against self antigens. This has meant that during vertebrate evolution the immune system has developed strategies to learn to recognise and respond to the antigenic universe both during ontogeny (somatic recognition strategy during life in individual animals) and phylogeny (a germ line strategy for antigen recognition). The founding concepts of modern immunology are based on the Clonal Selection Theory of Acquired Immunity of Sir MacFarlane Burnet (Burnet 1959). A good summary of the theory can be found in the *Scientific American* article by Ada and Nossal published in 1987.

### Darwinian Selection

Thus 'Darwinian Selection' principles are the first key component of the learning mechanism at the molecular, cell and whole animal levels. Clones of somatic B or T lymphocytes live or die depending on their recognition of 'self' or "non-self". In particular somatic mutant B cells (below) live or die depending on the binding affinity of their mutant cell surface antibody receptor for antigen. With respect to B cells The Clonal Selection Theory posits an array of B cells emerging early in ontogeny each displaying a different antigen receptor on their surface membrane *i.e.* "one cell makes one antibody". Early proof of this concept was provided by Sir Gustav Nossal (Ada & Nossal 1987). Of course the most immediate antigens an emerging B cell from the bone marrow would confront would be those proteins and carbohydrates (mainly in membranes but also in the fluid phase) presented by "self". Burnet postulated a clonal deletion or purging mechanism to rid the body of such potentially auto-reactive lymphocytes. And we now know after 50 years of research that a variety of mechanisms help to establish and maintain self tolerance based on Burnet's 'clonal deletion' concept – they include clonal abortion, clonal anergy, clonal suppression (via regulator T cells), V receptor editing and V gene replacement (Nemazee 2006; Chen *et al.* 1995).

All those other B cells that escape this process of course constitute the anti-nonself recognition repertoire from which foreign antigens select anti-nonself clones (see Figure 1) and many go on to become longer lived 'memory cells' for recall later by the same antigen. Similar, although different processes, of 'negative' and 'positive' selection, occurs for T cells as they mature in the thymus (Steele *et al.* 1993).

An important modification was made to Burnet's theory in the 1970s when Alistair Cunningham proposed that most antibody diversity actually appears *after* antigenic stimulation, that is during the course of the immune response (Cunningham 1977). We now recognise this process by the phenomenon of antigen driven "somatic hypermutation" of rearranged immunoglobulin V genes (VDJs) which underpins the phenomenon of the affinity maturation of antibodies during an immune response (Berek & Milstein 1987, 1988).

### Lamarckian Gene Feedback Loops

The other learning mechanisms contributing to the repertoire of antigen recognition in both ontogeny and phylogeny we propose invoke roles for reverse transcription and soma-to-germline gene feedback (Steele 1979; Steele *et al.* 1998; Blanden *et al.* 1998; Steele *et al.* 1998). These will be elaborated on further below. In short, such processes allow 'directional' fine tuning and maintenance of functional V gene repertoires both during life and over evolutionary time.

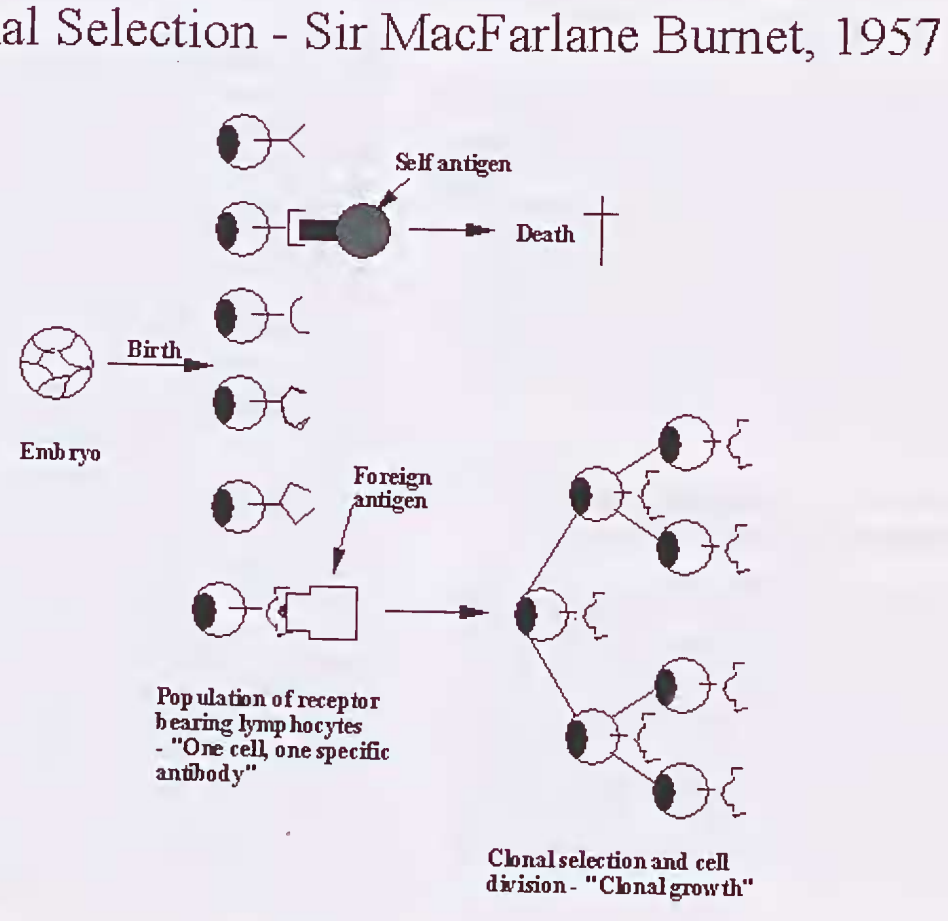


**Figure 1.** Clonal selection.

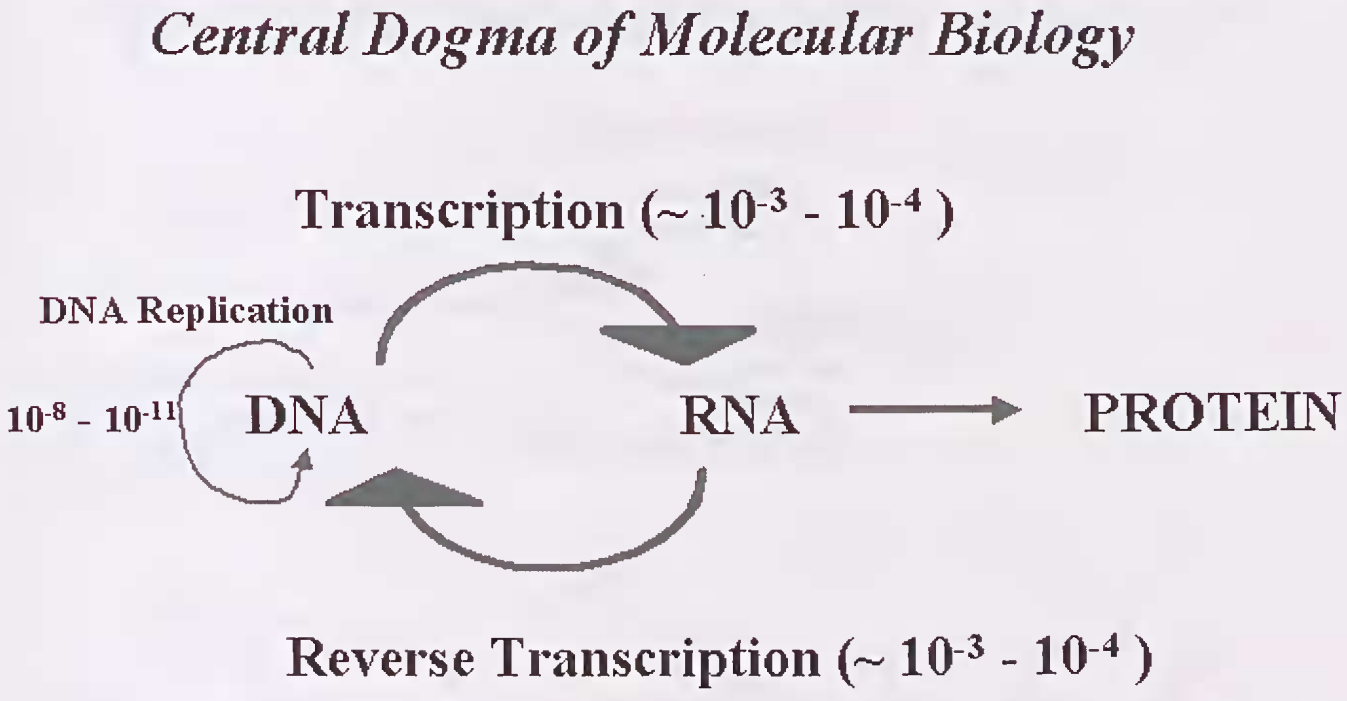


Figure 2. Central dogma of molecular biology: the frequencies represent point mutations per replication event per bp. Thus DNA replication is a relatively high fidelity copying process with errors occurring at a maximum rate of about one nucleotide substitution per 100 million bases replicated.

These feedback loops are enshrined in, or are modifications of, the Central Dogma of Molecular Biology (Figure 2) and Weismann's Barrier (Figure 3) – the first defines the rules for information flow at the molecular level and the second, a cellular theory, prohibits somatic genetic information(RNA/DNA) being fed-back to germline DNA in germ cells. This author believes that these two fundamental 'biological theorems' have been confused in the past, particularly in such widely read books by the philosopher and author Arthur Koestler (Koestler 1978). It was proposed that the Weismann Barrier can be selectively breached without violating what we know about molecular biological processes in general and be compatible with Darwinian natural selection principles – thus it could be imagined that somatic genetic information in the form of mRNA amplified in the soma could be available for transfer and integration into the genome of a germ cell and thus be inherited by progeny. This formed the basis of the Somatic Selection Hypothesis (Steele 1979) summarised schematically in Figure 4. For V genes in the immune system is was imagined that antigen-selected mutant VDJ genes could be selected by antigen and the mRNA taken up by a harmless endogenous retroviral vector (endogenous RNA virus, ERV) and delivered to germ cells where the RNA could be copied into DNA via a reverse transcriptase (thought to be provided primarily by the ERV). Independent support for this general scheme can be seen in the ground breaking work of Corrado Spadafora and colleagues initiated in the late 1980s which will be discussed below (reviewed in Smith & Spadafora 2005; Spadafora 2008).

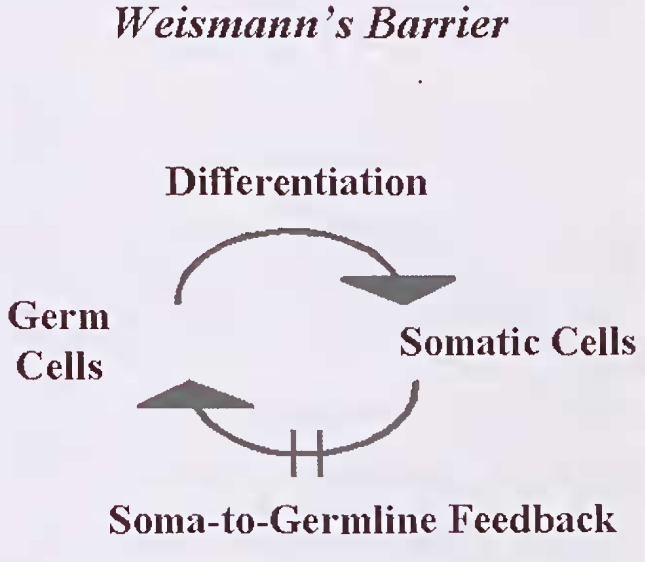


Figure 3. Weismann's Barrier.

The mechanism of somatic hypermutation appears to also require a RNA intermediate and thus a reverse transcriptase step to fix somatic RNA mutations in lymphocyte DNA (Steele & Pollard 1987; Steele *et al.* 1997; Steele 2009). We will discuss these data and analyses in more detail below.

## Research Strategy in Antibody Diversity

For over 25 years our research strategy has depended on two parallel research programs, a) an analysis of the mechanism of antigen-driven somatic hypermutation, and b) how this somatic genetic diversity might impact on the diversity and "genetic quality" of the germline V

# *Weismann's Barrier Selectively Permeable*

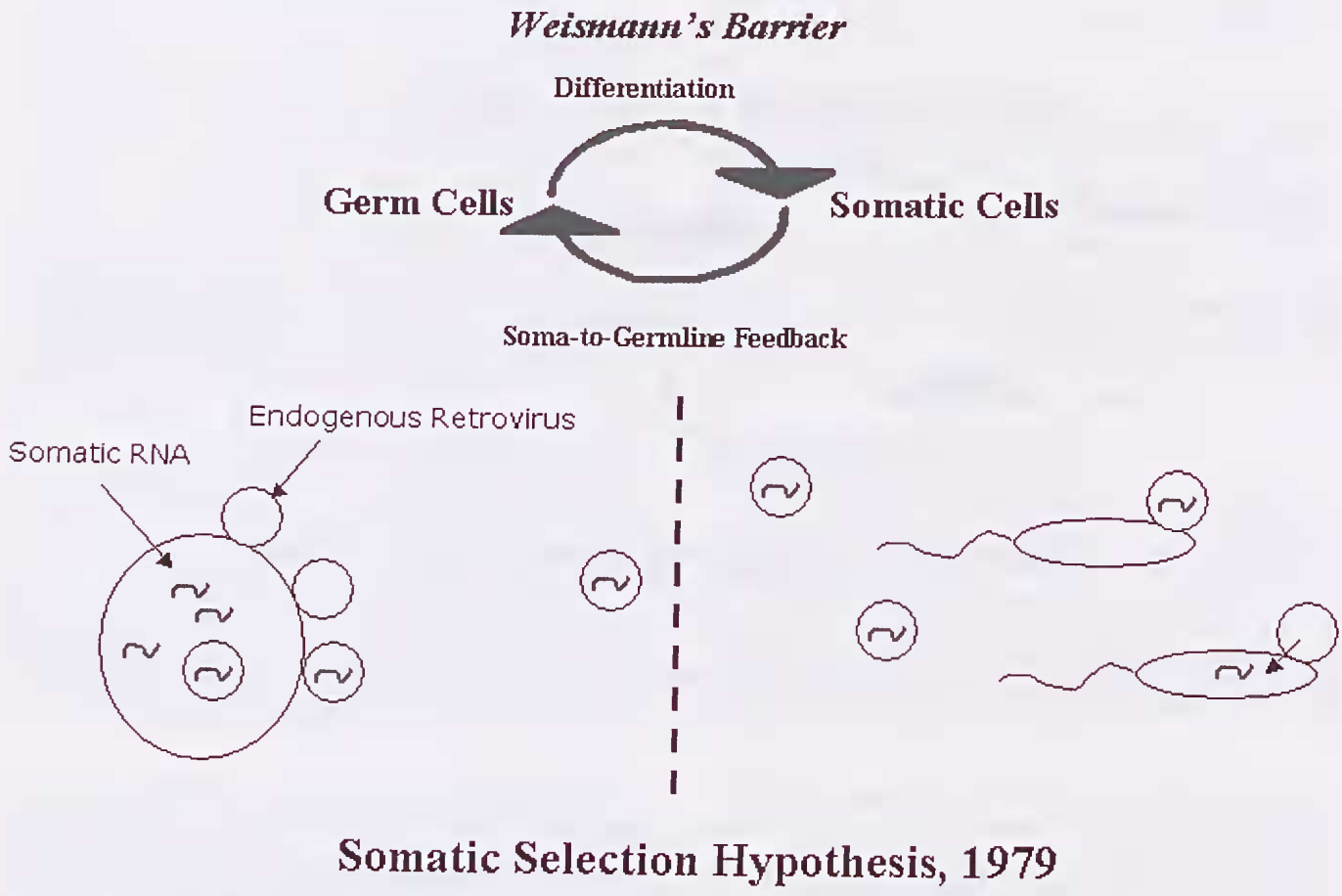


**Figure 4.** Somatic selection hypothesis, 1979.

gene arrays at the immunoglobulin locus in vertebrates. This work has allowed us to conclude that a significant portion of the somatic mutation and antigen-selection pattern in antibody variable genes (~80%) is indeed written into the germline V gene arrays at Ig loci.

### Rearrangement, Gene Expression and Somatic Hypermutation of VDJ genes

A critical factor in the analysis depends on what has been established about the germline and somatic expression of immunoglobulin genes – they have clear 'germline' and 'somatic' configurations (Figure 5 and see Honjo *et al.* 2004). These facts allow us to infer and deduce that genetic information has indeed flowed from the somatic compartment to the germ cell compartment over evolutionary time.

The figure shows a schematic outline of a mammalian immunoglobulin heavy chain in its germline configuration and its somatic configuration. The germline, or unrearranged DNA configuration, exists in germ cells and all non-lymphoid cells in the body (*e.g.* kidney cells, liver cells, *etc.*, Honjo *et al.* 2004). Thus on the left hand side (5' side) are the array of so-called 'V-elements' or 'V-segments', which would typically encode approximately 95 amino acids and the 100–200 V-elements are encoded in a span of chromosomal DNA of about 1 Mb in the human genome . This repertoire of unrearranged V-elements lies about 100Kb upstream (in the transcriptional sense) of very short genetic elements termed diversity (D) and joining (J) regions. There are 10–30 D regions and 4–5 J regions at typical mammalian IgH loci (together they would encode after VDJ assembly approximately 25 additional amino acids). Further downstream, encompassing about 10 Kb lie successively the intronic enhancer and nuclear Matrix Attachment Region (EiMAR) and then the Ig class switch region DNA repeat elements and then a series of constant region exons encompassing Ig heavy chain isotypes, mu (υ), delta (δ), the various γ chain subclasses (IgG1, IgG2a, IgG2b IgG3) and then the α chains for secretory IgA and ε chains for mast cell binding and allergy-activating IgE antibodies. Further downstream is the 3' enhancer region (Honjo *et al.* 2004).

The various Ig classes reflect the functional properties of the antibody once antigen has been bound by the antigen combining site. The functional properties would include Complement activation and thus opsonisation of foreign particles for phagocytosis by monocyte scavenger cells (*e.g.* macrophages). The antigen binding site is a heterodimer of a light (L) and heavy (H) chain so antigen binding and thus antigen-mediated selection can only

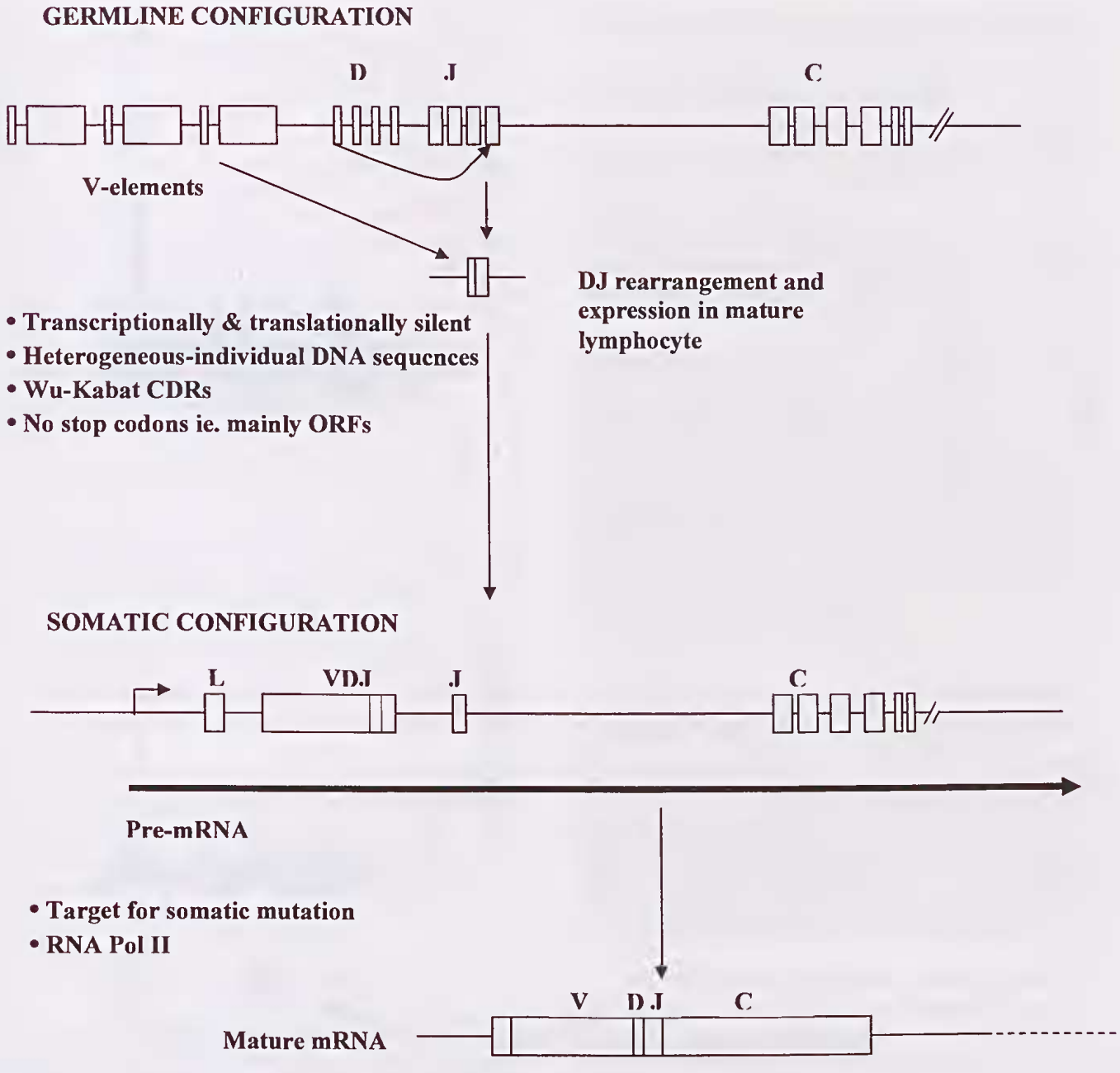


**Figure 5.** Rearrangement and immunoglobulin gene expression.

occur on a fully assembled Ig molecule or a B cell that displays such an antigen receptor on its surface membrane.

A key process not shown in Figure 5 is immunoglobulin class switching (CSR) whereby following cytokine signalling from other lymphocytes (T cells) and white cells, the B cell will switch from transcribing and assembling IgM heavy chains (υ chains), and reposition the productively assembled VDJ gene further downstream in front of one of the down stream Ig isotypes. This is a looping out DNA recombination mechanism such that the B cell retains the original selected VDJ but now has it joined to a different set of constant region exons (Honjo *et al.* 2004).

It is already clear that Ig loci display a degree of genetic complexity not observed in more straight forward single-copy house keeping or tissue-specific protein coding genes. T-cell receptor genes also display the same general genetic organisation and expression strategy (but typically do not normally somatically hypermutate their assembled VDJ genes). There is evidence that the protocadherin synaptic receptor genes in the central nervous system show a similar variable-to-constant rearranging strategy as seen for immunoglobulins and T-cell receptor genes but in these cases it is executed at the RNA level by an alternative splicing mechanism (for a mini-review see Chess 2005).

Germline V-elements *per se* are never transcribed into RNA for inclusion in a mRNA prior to translation into Ig proteins. In this sense they are transcriptionally and translationally silent. As such V-elements or their products are never the direct targets of antigen-binding selection. This type of antigen-mediated somatic selection is only directed to a fully assembled VDJ gene in the

context of a light chain VJ gene co-expressed in the same B cell (and thus clonally selected).

VDJs are therefore the substrates for both RNA polymerase II transcription and somatic hypermutation. It is important to note that additional mutational errors are introduced by the DNA rearrangement process at the V-D and D-J borders (termed 'junctional diversity') and because the process is stochastic only a minor portion (about 10% of all rearrangements of IgH chains) are 'productive' *i.e.* in the correct translational reading frame (Honjo *et al.* 2004). This critical point will be discussed later in the context of fused 'VD' pseudogenes at chicken IgH loci (Rothenfluh *et al.* 1995).

Thus a germline repertoire of 100 functional VH elements, 20 D and 5 J regions can theoretically encode 100 x 20 x 5 or 10,000 VDJ regions; similarly 100 functional VL elements and 5 J regions would encode 500 VL. Together there is a potential combinatorial germline repertoire of perhaps 10,000 x 500 or 5 x $10^6$ unique antibody specificities. Following antigenic stimulation and somatic hypermutation this potential repertoire could perhaps increase another order of magnitude or two (Berek & Milstein 1987,1988). In reality ongoing antigen selection sifts and focuses the response for higher affinity antibodies so the full potential is unlikely to be realised in any individual (and of course the upper limit would be set by the total number of B cells generated from the bone marrow at any given point in time).

To summarise, genetic information in the form of unrearranged V-elements is never subject to direct antigen binding selection on the intact antibody or Ig receptor bearing cell. In contrast B cells expressing fully rearranged VDJ (heavy chain) and VJ (light chain) genes are subject to direct antigen-binding selection. It is this crucial distinction that demarcates the germline from the somatic configuration and thus allows deductions on the origin of highly non-random DNA sequence patterns.

### The Germinal Centre and Affinity-Based Selection

A naïve B cell in the periphery can be selected by antigen to immediately secrete its encoded antibody or it can migrate to the primary follicle in lymphoid tissue to become a founder B cell in a Germinal Centre (termed GC). One or just a few B cells locate in a follicle and they multiply to form small colonies of 10,000 to 20,000 cells. Due to antigen-binding competition between pre-existing low affinity antibody and antigen-antibody complexes displayed on follicular dendritic cells within the GC, only the mutated B cells displaying viable high affinity antibodies survive – the rest die by the programmed cell death process called apoptosis (≥90% of all B cells in a Germinal Centre die there). In this way the mutated B cell survivors become antibody secreting cells and memory cells and they bear the signature of non-random DNA sequence modifications typical of intense selection. That is, point mutations in the VDJ accrue in those regions termed CDR or Complementary Determining Regions, which encode the amino acids that make direct contact with the molecular shapes of the antigen (typically protein, carbohydrate). Typical 'Wu-Kabat' plots of this non-random variability are shown in Figure 6 for 30 somatically mutated derivatives of the rearranged VH186.2 gene in mice (Steele *et al.* 1993). All of these features of the Germinal Centre reaction have been covered at length elsewhere eg. in *Lamarck's Signature* (Steele *et al.* 1998) or can be found in more specialised publications (*e.g.* MacLennan 1994).

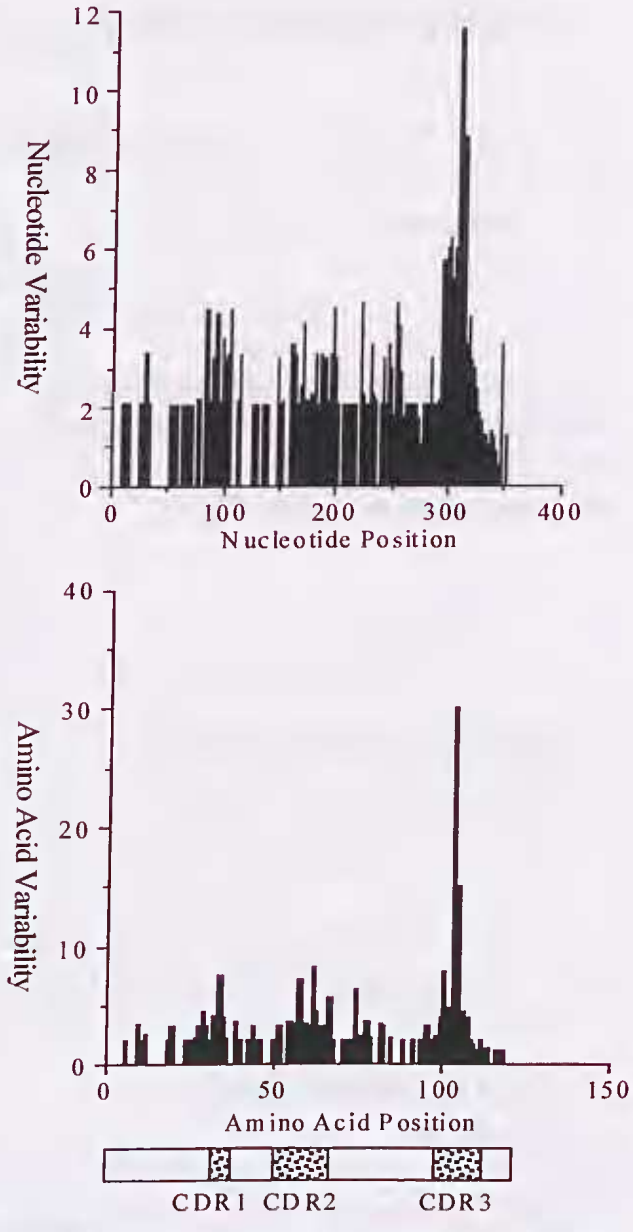


**Figure 6.** Wu-Kabat variability plots. Thirty somatically mutated derivatives of the mouse VH-1186.2 heavy chain variable gene assembled as V(D) is in mature anti-NP antibodies. Variability = Number difference at a position / Frequency of most common at that position. VDJ DNA sequences at top; translated protein sequences at bottom. Adapted from Steele *et al.* (1993).

### The Mechanism of Somatic Hypermutation

The dominant current model of somatic hypermutation, "The DNA Deamination" model is DNA-based. *i.e.* all the mutational events occur directly at the DNA level (Di Noia & Neuberger 2007; Teng & Papvasiliou 2007). The main first step entails deamination of Cytosine to Uracil by the enzyme activation-induced cytidine deaminase (AID) which targets Cytosines in the context of WRCY hot spots (W = A or T, R = A or G purines, and Y = C or T pyrimidines). The resulting C-to-U lesions in DNA are either repaired by a base excision DNA repair pathway (involving uracil

DNA glycosylase, UNG), or if not repaired, replicated over to produce C-to-T mutations. If repaired by UNG the resulting abasic site can be transformed into a nick in the DNA by an endonuclease termed apurinic apyrimidinic endonuclease (APE). Alternatively the G:U mispairs attract the mismatch repair heterodimer MSH2-MSH6 which also recruits the error-prone translesion DNA polymerase-η (eta) which introduces mutations in the repaired patch by targeting A:T base pairs at WA-sites (where W = A or T). These series of steps are very similar to the V targeted-nicking and error-prone repair model of somatic mutation of IgV genes first advanced by Brenner and Milstein in 1966.

However three sets of recent observations by our group are not easily reconciled with the standard model but are consistent with, or predicted, by the reverse transcriptase model (Steele 2009):

a) DNA polymerase-η the accepted and *sole* A:T mutator in SHM is an efficient reverse transcriptase *in vitro* (Franklin *et al.* 2004);

b) The RNA editing signature of ADAR1-mediated A-to-I deamination, as instanced by the elevated A-to-G mutations, is embedded within the SHM pattern (Steele *et al.* 2006); and

c) The AID deaminase-linked RNA polymerase II RNA mutation signature, as instanced by elevated mutations at G sites (particularly G-to-C and G-to-A) is embedded within the SHM pattern (Steele 2009).

This work has led us to conclude that the weight of evidence now favours "The Reverse Transcriptase Model first advanced by Steele and Pollard in 1987. Thus somatic hypermutation in B lymphocytes involves:

a) Direct DNA deamination (C-to-U, thus giving rise to C-to-T and G-to-A mutations);

b) RNA Pol II copying deaminated DNA templates carrying U and abasic site lesions generates mutated mRNA (giving rise to G-to-C and G-to-A strand biased mutation signatures);

c) RNA deamination (editing) of mRNA causing Adenosine-to-Inosine mutations via ADAR1 deaminase (thus causing the A-to-G strand biased mutation signature); and

d) Error-prone reverse transcription by DNA Pol-η to fix the RNA mutations in B lymphocyte DNA and create further strand biased mutations *viz.* the transversions A-to-C and A-to-T.

To summarize, during somatic hypermutation both direct DNA mutations and a variety of RNA mutations are copied back into DNA (Steele 2009). As we have pointed out earlier (Steele *et al.* 1997) a role for a *cellular* reverse transcriptase such as DNA polymerase-η acting in its reverse transcriptase mode should no longer be a heretical concept given that "telomerase", a ribonucleic acid-protein particle has as its core function the capacity to copy the RNA repeat into a DNA repeat – a critical step in synthesis of the telomere multiple repeat chromosomal cap. That is "telomerase" itself is a cellular reverse transcriptase (Blackburn 1992).

## Evolution of the Germline V-Segment Repertoire and Soma-to-Germline Feedback

Is there any evidence that somatically mutated variable genes can be fed back to the germline V-segment repertoire? It is possible to answer this question in the affirmative because germline V-segments can *never* be the targets of somatic hypermutation nor direct antigen-mediated selection at the protein level *i.e.* on an intact Ig antigen receptor on a mature B cell surface. This fact is

**1. Pristine condition ("open reading frames" or minor defects)**

**2. Non-random mutation patterns (= Ag mediated somatic selection)**

**3. Insertion-deletions display a highly non-random distribution :**
*** protein coding regions : triplet codons or multiples thereof**
*** non-coding regions: variable length**

**4. Chicken V pseudogenes show non-random mutation patterns and features of somatic V->D->J rearrangement signature**

**5. Flanking and coding regions are evolving independently**

**6. Major sites of genetic recombination coincide with predicted sites of soma-to-germline retrotransposition**

Figure 7. Key features of the repertoire Germline V Segments (Blanden *et al.* 1998).

often not addressed nor appreciated by those in the field speculating on the evolution of the V-segment repertoire. As we have documented elsewhere the germline V repertoires of families of similar V segments display all the hallmarks of strong somatic mutation and antigen-mediated selection i.e. a significant portion of the somatically mutated VDJ repertoire generated during evolutionary time in vertebrates has been fed back to the germline most likely by a reverse transcriptase intermediate step and targeted to unrearranged V genes by homologous recombination (approx ≥80% of the assembled VDJ gene is comprised of the V element). Figure 7 summarises the main findings of this work which has been extensively published in refereed literature (Rothenfluh *et al.* 1994; Rothenfluh *et al.* 1995; Weiller *et al.* 1998; Blanden *et al.* 1998; Zylstra *et al.* 2003).

In short, a highly non-random somatic mutation and selection signature dominates the DNA sequence pattern of families of similar V genes arrayed, usually in tandem, in the vertebrate germline.

As suggested above these are 'subtle' somatic mutation signatures – they are of a different class from the more obvious retro-sequence impact events that dominate vertebrate and mammalian genomes. They argue for a requirement for innovative ways of interpreting the DNA landscape of the genome. Thus sites of hyper-recombination near RNA splice site borders in the L-V intron make sense in this model (Weiller *et al.* 1998) as do the strange features of chicken VH pseudogenes which all have fused 'D' bits in the correct reading frame (Rothenfluh *et al.* 1995; Ota & Nei 1995).

From our research thus far we therefore conclude:

1. Somatic hypermutation of antibody V genes operates by direct DNA and RNA base modifications coupled to reverse transcription and integration of mutated cDNA retrotranscripts back into chromosomal DNA *within* a B lymphocyte.
2. Over evolutionary time somatically mutated and selected ("successful") V sequences from B lymphocytes have undergone homologous recombination into germline DNA, thus contributing to germline diversity *and* the maintenance of a *functional* germline V gene repertoire.

## Evidence from other systems

Is Soma-to-Germline feedback a general phenomenon in complex biological systems? We address this by briefly reviewing the work of other groups.

### Corrado Spadafora

From about the late 1980s to the present Corrado Spadafora and colleagues in Rome have published a series of important papers clearly showing that mammalian spermatozoa can take up foreign nucleic acid molecules and express the genetic information in progeny organisms. In particular mouse spermatozoa *in vitro* can absorb both foreign DNA and RNA, and if the latter then a LINE-1-derived reverse transcription step will be executed copying the RNA into DNA. In a small number of cases (≤ 10% ) the DNA sequences are integrated into the germline genome. In the majority of cases the sperm-absorbed DNA/RNA exists as extrachromosomal episomes which replicate along with the host somatic cells during development displaying mosaic tissue expression (see reviews in Smith & Spadafora 2005; Spadafora 2008). This work clearly shows there is no physical barrier to uptake of DNA or RNA, although there maybe clear developmental stages in spermatogenesis when spermatozoa are susceptible to foreign nucleic acid uptake (Zoraqi & Spadafora 1997).

### Patrick Fogarty

Using an innovative technique based on P-elements and delivering DNA transgenes intravenously in simple vesicles, Fogarty has shown that 50% of progeny from such male mice inherit the gene sequence (Fogarty 2002). The critical integration event requires a transposase. This works suggests that non-cellular DNA can readily transverse the testes tissue barriers, that normally quarantine the production of sperm, be integrated into the germline and be transmitted to progeny.

### Minoo Rassoulzadegan

The group of Minoo Rassoulzadegan has shown that mature sperm carry more than just a compact haploid DNA nucleus (Rassoulzadegan *et al.* 2006). Thus sperm heads contain gene-specific regulatory RNAs (miRNA) which at fertilisation can have profound genetic effects in progeny. The phenomenon described involves an allele-specific "paramutation" effect of the *Kit* locus (important in mammalian development) yet the implications of the finding are far reaching. Certainly the effects can be transmitted to an additional breeding generation. The mechanism is unclear given that animal miRNA systems are not thought to amplify their miRNA precursors by a double stranded RNA polymerase as in plants. Perhaps the transgenerational effects are based on long lived RNA molecules? However the whole phenomenon raises the possibility of germline fixation of such epigenetic intermediary effects via a reverse transcription step at some level of the RNA regulatory process.

### John Mattick

Since the mid 1990s John Mattick and colleagues have been documenting the extent and importance of the RNA regulatory networks of what we now call the extended 'transcriptome" (Mattick 2007; Mattick *et al.* 2009). Thus non-protein coding regions produce ncRNAs which are regulatory in nature, regulating gene-specific expression of protein coding genes. Only about 2% of the entire mammalian 'transcriptome' codes for proteins – the rest (>98% genome) are involved in specific gene regulation in a multilayered complex best described as the "RNA regulatory universe". In more recent papers Mattick concedes the necessity for some form of soma-to-germline feedback to be operative (Mattick 2009) to ensure that selected genetic changes at this level contribute to the evolution of complex systems, particularly the brain and central nervous system (Mattick & Mehler 2008).

### Lars Holmgren

For a number of years Holmgren and colleagues have studied the potential genetic consequences in metastases of horizontal transfer of tumor genes via dispersal and

uptake of apoptotic bodies (Bersmedh *et al.* 2001; Ehnfors *et al.* 2009). These studies clearly show that if the laterally spreading DNA confers a selective advantage on the recipient cell then integration of the DNA is manifest and the DNA sequences propagated to progeny cells. We speculated on this type of somatic gene transfer in the late 1970s (Steele 1979).

### Genetic Cargo of Sperm?

This is an appropriate question given the interesting findings of Spadafora and Rassoulzadegan we have just discussed. Indeed when one considers an ovum at about the time of natural fertilization there are more questions than answers raised by the phenomenon (Figure 8).

Have we become so used to an image like this that we have forgotten how truly amazing it really is. Apparently one sperm succeeds in the race to fertilisation (one possibility is arrowed). Given conventional wisdom all the other attached sperm have no further say in the genetic outcome. Apart from the non-Mendelian routes of genetic transfer in the experiments of Spadafora and Rassoulzadegan we have known for many years (*e.g.* Keissling *et al.* 1987) that sperm heads have clusters of attached endogenous retroviruses – to what end one might ask? Moreover, ERV concentrations are very high in seminal fluid ($\geq 10^{11}$ per ml) and ERVs are emitted in copious quantities from activated lymphocytes (they also are prominent in Germinal Centres following immunisation). ERVs have been observed coating the female placenta (see Rothenfluh 1995 for more references of this type). Again the question arises – to what biological purpose should cells of the immune system and reproductive tissue be so predominantly associated with either ERV production or unexpected ERV tissue localisation?

## Concluding Remarks

We conclude that both Darwinian antigen-binding selection and Lamarckian soma-to-germline feedback play key roles in the evolution of antibody variable genes. There is also evidence supporting the view that reverse transcription is central to a better understanding of the somatic and germline evolution of these genes. The work of a number of groups suggest that the ease of gene movement between cells, whether they be germline or somatic, suggests that soma-to-germline feedback is likely to be general in complex biological systems and contributes to genome diversity. Thus Weismann's Barrier is viewed as being *selectively permeable* to somatic genetic information provided it is beneficial to both parent and progeny organisms (Figure 9). Acquired somatic genetic information ("experience") may therefore not be lost with the death of the individual but be propagated to progeny who would then be selected in a Darwinian manner for fitness.

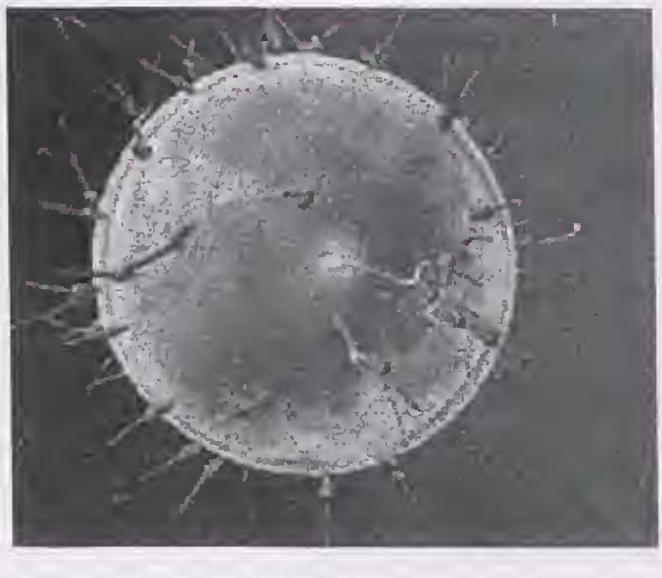

Figure 8. Sperm fertilising ovum.

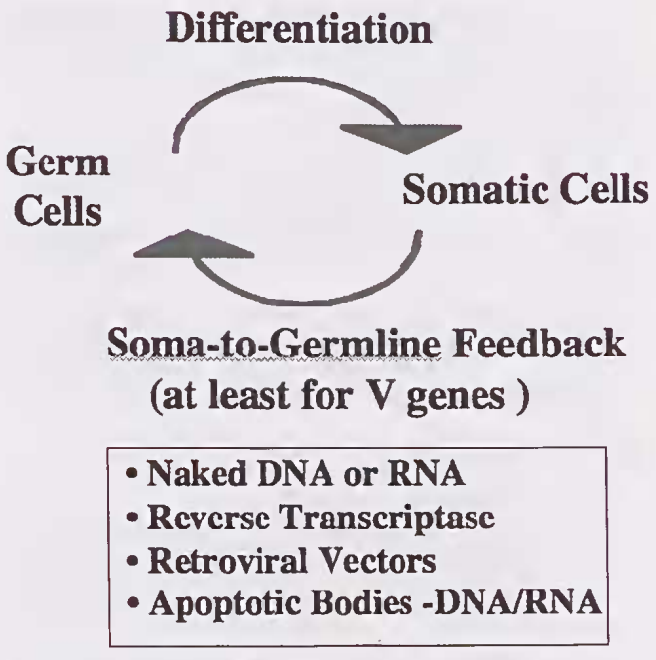


Figure 9. Weismann's Barrier selectively permeable.

## References

Ada G L, Nossal G 1987 The Clonal Selection Theory. Scientific American 255: 50–57.

Bergsmedh A, Szeles A, Henriksson M, Bratt A, Folkman M J, Spetz A L & Holmgren L 2001 Horizontal transfer of oncogenes by uptake of apoptotic bodies. Proceedings of the National Academy of Sciences of the United States of America 98: 6407–6411

Blanden R V, Rothenfluh H S, Zylstra P, Weiller G F & Steele E J 1998 The signature of somatic hypermutation appears to be written into the germline IgV segment repertoire. Immunological Reviews 162: 117–132.

Berek C & Milsten C 1987 Mutation shift and repertoire drift in th maturation of the immune response. Immunological Reviews 96: 23–41.

Berek C & Milsten C 1988 The dynamic nature of the antibody repertoire. Immunological Reviews 105: 5–26.

Brenner S & Milstein C 1966 Origin of antibody variation. Nature 211: 242–243.

Burnet F M 1959 The Clonal Selection Theory of Acquired Immunity. Ambridge University Press, London and New York.

Chen C, Nagy Z, Prak E L & Weigert M 1995 Immunoglobulin heavy chain gene replacement: a mechanism of receptor editing. Immunity 3: 747–755.

Chess A 2005 Monoallelic expression of protocadherin genes. Nature Genetics 37:120–121.

Cunningham A J 1977 Evolution in microcosm: The rapid somatic diversification of lymphocytes. Cold Spring Harbor Symposium. Quantitative Biology 41: 761–770.

Di Noia J M & Neuberger M S 2007 Molecular mechanisms of antibody somatic hypermutation. Annual Review of Biochemistry 76: 1–22.

Ehnfors J, Kost-Alimova M, Persson N L, Bergsmedh A, Castro J, Levchenko-Tegnebratt T, Yang L, Panaretakis T & Holmgren L 2009 Horizontal transfer of tumor DNA to endothelial cells in vivo. Cell Death and Differentiation 16: 749–757.

Fogarty P 2002 Optimizing the production of animal models for target and lead validation. TARGETS 1: 109–116 (Journal now part of "Drug Discovery Today")

Franklin A, Milburn P J, Blanden R V & Steele E J 2004 Human DNA polymerase-η, an A-T mutator in somatic hypermutation of rearranged immunoglobulin genes, is a reverse transcriptase. Immunology and Cell Biology 82: 219–225.

Honjo T, Alt F W & Neuberger M 2004 The Molecular Biology of B Cells. Elsevier Academic Press, London and San Diego.

Keissling A A, Crowell R C & Connell R S 1987 Sperm-assocaited retroviruses in the mouse epididymis. Proceedings of the National Academy of Sciences of the United States of America 84: 8667–8671

Koestler A 1978 Janus: A summing up. Hutchison, London.

MacLennan I C M 1994 Germinal Centers. Annual Review of Immunology 12:117–139.

Mattick J S 2007 A new paradigm for developmental biology. Journal of Experimental Biology 210: 1526–1547.

Mattick J S 2009 Has evolution learnt how to learn? EMBO Reports 10: 665

Mattick J S, Maral P P, Dinger M E, Mercer T R & Mehler M F2009 RNA regulation of epigenetic processes. Bioessays 31 51–59.

Mattick J S & Mehler M F 2008 RNA editing, DNA recoding and the evolution of human cognition. Trends in Neurosciences 31: 227–233.

Nemazee D 2006 Receptor editing in lymphocyte development and central tolerance. Nature Reviews Immunology 6: 728–740.

Ohta T & Nei M 1995 Evolution of immunoglobulin VH pseodogenes in chickens. Molecular Biology andEvolution 12: 94–102.

Rassoulzadegan M, Grandjean V, Gounon P, Vincent S, Gillot I & Cuzon F 2006 RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. Nature 441: 469–474.

Rothenfluh H S 1995 Hypothesis: A memory lymphocyte-specific soma-to-germline genetic feedback loop. Immunology and Cell Biology 73: 174–180.

Rothenfluh H S, Gibbs A J, Blanden R V & Steele E J 1994 Analysis of patterns of DNA sequence variation in flanking and coding regions of murine germline immunoglobulin variable genes – evolutionary implications. Proceedings of the National Academy of Sciences of the United States of America USA 91: 12163–12167.

Rothenfluh H S, Blanden R V & Steele E J 1995 Evolution of V Genes : DNA sequence structure of functional germ-line genes and pseudogenes. Immunogenetics 42: 159–171.

Smith K & Spadafora C 2005 Sperm-mediated gene transfer: Implications and applications. BioEssays 27: 551–562.

Spadafora C 2008 Sperm-mediated "reverse" gene transfer: a role of reverse transcriptase in the generation of new genetic information. Human Reproduction 23: 735–740.

Steele E J 1979 Somatic Selection and Adaptive Evolution : On the Inheritance of Acquired Characters. Williams-Wallace, Toronto. (2nd Edition, University of Chicago Press, Chicago, 1981).

Steele E J & Pollard J W 1987 Hypothesis : Somatic Hypermutation by gene conversion via the error prone DNA -to- RNA-to- DNA information loop. Molecular Immunology 24: 667–673.

Steele E J, Rothenfluh H S, Ada G L & Blanden R V 1993 Affinity maturation of lymphocyte receptors and positive selection of T cells in the thymus. Immunological Reviews 135: 5–49.

Steele E J, Rothenfluh H S & Blanden R V 1997 Mechanism of antigen-driven somatic hypermutation of rearranged immunoglobulin V(D)J genes in the mouse. Immunology and Cell Biology 75: 82–95.

Steele E J, Lindley R A & Blanden R V 1998 Lamarck's Signature : How retrogenes are changing Darwin's natural selection paradigm. Allen & Unwin, Frontiers of Science: Series Editor Paul Davies , Sydney, Australia, 1998. (In the USA published by Addison-Wesley-Longman under Perseus Book imprint, Reading, MT, 1998).

Teng G & Papavasiliou F N 2007 Immunoglobulin somatic hypermutation. Annual Review of Genetics 41:107–120.

Weiller G F, Rothenfluh H S, Zylstra P, Gay L, Averdunk H, Steele E J & Blanden *R V* 1998 Recombination signature of germline immunoglobulin variable genes. Immunology and Cell Biology 76: 179–185.

Zoraqi G & Spadafora C 1997 Integration of foreign DNA sequences into mouse sperm genome. DNA and Cell Biology 16: 291–300.

Zylstra P, Franklin A, Hassan K A S, Powell K L, Steele E J & Blanden R V 2003 Molecular evolution of VHGAM3.8-related germline genes isolated from DBA, BALB/c, 129 and C57BL mouse strains and sublines. Immunogenetics. 55: 182–188.